DOCUMENT RESUME

ED 321 579 FL 018 717

AUTHOR Djiwandono, M. Soenardi
TITLE Developing Indonesian Language Tests for College Graduates.
PUB DATE Apr 90
NOTE 11p.; Paper presented at the Annual Regional Seminar of the Regional Language Centre (Singapore, April 9-12, 1990).
PUB TYPE Reports - Descriptive (141) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Cloze Procedure; *College Faculty; Comparative Analysis; Foreign Countries; Graduate Students; Grammar; Higher Education; *Indonesian; *Language Tests; Multiple Choice Tests; Official Languages; *Test Construction; Test Validity; Writing Tests
IDENTIFIERS Indonesia

ABSTRACT

In Indonesia, Bahasa Indonesian (BI) is the designated national and official language. However, deficiencies in Indonesian proficiency are found in a wide range of individuals. A test battery to measure proficiency level was developed, consisting of a writing test, a grammar test, and a cloze test. The writing test was an essay, in which five components were assessed, namely, content, organization, vocabulary, language, and mechanics. The grammar test was a 50-item multiple-choice instrument measuring practical grammatical usage. The cloze test, developed as a possible alternative to the others, had a multiple-choice format. Three brief passages had words deleted at regular intervals. Subjects tested were 103 graduate students and 10 college lecturers. On the writing test, the graduate students' scores ranged from "poor" to "good," while the lecturers' scores were mostly in the "good" range. The test was found to be reliable and valid. Grammar test results showed a wide range of scores, with similar results from graduate students and lecturers. Areas for test improvement were identified. Cloze test scores, while higher than grammar test scores, showed similar patterns. Comparison of the results revealed the lowest correlation between writing and grammar tests, and the highest correlation between grammar and cloze tests. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED321579

Developing Indonesian Language Tests for College Graduates

M. Soenardi Djiwandono
IKIP Malang, Indonesia

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
M. Djiwandono
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Background

In Indonesia, Bahasa Indonesia (BI) or the Indonesian language, is designated as the national language that is used as part of national identity and a means of communication among a great number of ethnic groups that make up the Indonesian nation. It is also the official language for state and formal affairs in the country as well as medium of communication in the development of education, science and technology, and culture (PPPB, 1977). In such an important position, good mastery and use of BI is a crucial necessity, especially among those whose responsibility requires a frequent use of the language to express themselves and communicate with the people, orally or in written form. They include not only government officials, public figures, authors and journalists, but also scholars and intellectuals, and even students, especially those of advanced levels of education.

Unfortunately at the present stage of its application as the national and official language, the use of BI is still marked with deficiencies of various kinds, observable in the language of the mass of the people and even those with better educational background. This phenomenon has been frequently pointed out by different observers at different occasions and has been the subject of numerous studies conducted to identify the deficiencies more accurately. For that purpose articles have been written, and research results reported, indicating that the inadequate mastery of BI is in fact widespread in terms of kinds of language skills and components, and also in terms of levels of education of the users. To that effect Badudu (1975, 1985) and Samsuri (1985), for example, express their regrets about the poor mastery of BI of the Indonesian people in general, including high school and college students, and even university lecturers (Sadtono, 1976). Meanwhile Suharianto (1977) found out that the defiencies in the use of BI can be found in almost every aspect of langauge use: the choice of words and phrases, arrangement of words, sentence patterns, the use of prepositions etc. Inadequate reading ability was revealed by Sunaryono Basuki et al (1986), as was writing ability of college students by Syafi'ie (1984), all of which add up to the fact, however unpleasant, that the mastery of the national language of the country by its citizens, is not up

FL018717

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

to a satisfactory level.

However, the numerous statements and research conclusions underlining the lack of competence in the use of BI have mostly been based on general observations and impressions of the failure to use the appropriate forms of BI. They have usually been in the form of analysis of the kinds and frequency of mistakes and inaccuracies that can be detected in the use of the language. Apparently, not much has been done to try to qualify such statements more precisely by describing the exact levels of the mastery of BI. For that purpose no proper instruments are readily available and efforts are yet to be made to produce standardized or well-constructed tests. This paper presents the summary of such an initial attempt conducted recently at the the graduate school of IKIP Malang in Indonesia (Djiwandono, 1990).

The tests to be developed

As an initial attempt to develop a test battery by which the level of mastery of BI can be measured more objectively the study was limited to the construction of tests for the educated people represented by students of some graduate schools. In their daily work, academic or otherwise, these educated people very often have to express themselves in BI, especially in writing, in a clear and orderly manner in order to communicate effectively. This limited the study further to the development of test of writing for the educated people as the first part of the battery that was to be given priority.

In addition to the writing test as the main emphasis of the present test battery development, attempts were also made to explore the possibility and feasibility of developing other kinds of tests so that if necessary one's competence in BI could be evaluated in different ways, using different tests. The other kinds of tests to be developed in the study included a test of grammar and cloze test. The inclusion of grammar test was due to conclusions and opinions expressed in various publications, implying the inadequate mastery of Indonesian grammar among its users. A well-constructed test of grammar may help verify the impression more objectively. Meanwhile a cloze test was also included in the development of the battery as further effort to explore its suitability, and effectiveness, in testing competence in BI.

In developing the test attempts have been made to keep the emphasis on testing <u>ability to use</u> the language and to avoid the testing of <u>knowledge about</u> BI. This was an important factor in the development of tests of BI, considering the general tendency and practice common among many teachers of BI to give more emphasis on theory and knowledge about language in teaching, and consequently also in testing, than the more practical aspect of usage. In that line of thinking the three kinds of tests: writing, grammar

and cloze, were all constructed strictly for testing the ability to write, to use grammar and to use the language.

Construction of tests

Writing test

The test of writing as the main part of the battery to be developed was intended as a means of assessing writing ability directly from a writing performance rather than indirectly through the assessment of its subskills tested separately. This was done by assigning testees, who were students of some graduate schools, to produce a piece of writing in BI in accordance with some preset requirements and directions. These included the title ("The issue of quality vs. quantity in Indonesian education"), a brief description of the background information, the length of essay (one page of folio-sized, lined writing paper), and the time allocation (30 minutes).

Evaluation of the essay was done holistically by basically applying "ESL Composition Profile" (Jacobs et al, 1981). The profile contains five component scales, "each focusing on an important aspect of composition and weighted according to its estimated significance for effective written communication" (p.91). The five components include content (30 points), organization (20 points), vocabulary (20 points), language (25 points) and mechanics (5 point). The total weight for each component is further broken down into numerical ranges that correspond to four mastery levels of excellent to very good, good to average, fair to poor, and very poor. The range of total scores thus extends from a minimum of 37 to a maximum of 100.

In the application of the profile to the development of writing test of BI, the component of Language was interpreted as Tatabahasa or grammar. This was found to be more relevant to the Indonesian setting where use of grammar of BI was still a problem. Besides the specification of Language in the profile also implies grammar. In addition to that an exchange of position in the order of components was made between Language (now Tatabahasa or grammar) and Vocabulary. In this rearrangement Tatabahasa, which is considered to be more essential, precedes Vocabulary.

Grammar test

The grammar test was constructed as a 50-item multiple-choice test. Each item required the testees to rely on their ability to use the correct, acceptable grammatical point rather than their theoretical knowledge about it. The selection of the grammatical points was based on the standard

grammar of BI as it is described in "Tatabahasa Baku BI" (Departemen P dan K, 1988). The points were derived from ten grammatical categories comprising verbs, nominatives, adjectives, adverbs, particles, sentence components, clauses, simple sentences, sentence modifiers, and complex sentences. For each category 5 items were prepared.

The test was written in multiple-choice format of two different forms: a multiple-choice completio. for the first section of the test (Section A), and a basically simple multiple-choice for the second (Section B).

The first form was used for the 25 items of section A, dealing with the correct use of the first five, single-word grammatical categories: verbs, nominatives, adjectives, adverbs and particles. Each item in this first section, as well as all other items in the grammar test, was originally a grammatical sentence taken from a text of a well-edited source, "Ensiklopedi Nasional Indonesia" or ENI (1988). Each sentence was selected so as to get an example of a certain grammatical point to be tested in actual use. A part of the sentence containing the grammatical point was then used as the correct option of the test item. Distractors for this option were obtained from real incorrect answers supplied by participants of a preliminary test conducted in the process in the form of a completion test where the key was simply deleted from the original sentence. The following is an example of an item in section A:

> Dengan curah hujan yang cukup tinggi, sekitar 3.200 mm setahun, dan ... ... , Kabupaten Agam tergolong subur.
>
> A. banyaknya hutan lebat
> B. banyaknya sungai kecil
> C. tingginya kelembaban tanah
> D. suhu udara normal
> E. dikelilingi sejumlah gunung

The second section of the grammar test was meant for items on the remaining five grammatical categories related more to sentence construction than individual words. Each item was a straightforward multiple-choice item containing five options including the key, or rather the expected response. Different from the keys of the items in the first section, the expected responses in the second section of the test were in fact words that made the sentences ungrammatical. Ability to identify these words was to be understood as an indication of mastery of grammar in use. Therefore an item in this section was a sentence that was completely quoted from the reading source one part of which was key to the grammaticalness of the sentence. This part of the sentence was to be replaced by another word that was often wrongly and ungrammatically used by many people. For a

more realistic sample of this kind of misused word, the same sentence had been used earlier in a preliminary test with the word deleted to leave a blank to be filled in. One of the words supplied for the blank was then chosen as the expected response, representing an element that the sentence ungrammatical. Meanwhile the distractors were in fact parts of the sentence that did not affect the grammaticalness of the sentence, but people with inadequate mastery of grammar might think otherwise. These were obtained in a similar manner as the expected response during the preliminary development of the test. An example of test item in the second section would read as follows:

Kewajiban (A) untuk mengumumkan (B) harta kebanyakan diatur oleh (C) sebuah (D) undang-undang "Public Law 95-521" yang juga (E) disebut "Ethics in Government Act of 1978"

(the correct choice is C, oleh which should actually be dalam).

Cloze test

The cloze test was included in the test battery to be developed at this initial stage as a possible alternative to the other two tests. It was constructed in multiple-choice format making the test a multiple-choice cloze.

To get 50 test items, three short passages were obtained from the same source that was also used for the grammar test. They were texts on subjects that were considered to be within the common knowledge of the educated people in general. Deletion was made every 10^{th}. word in the first text, and every 5^{th} word for the other two texts. Distractors for the test items were obtained in the same manner as the grammar test by administering the test in the completion form at the preliminary stage.

Test validation

At this initial stage an account of the suitability of the tests was based primarily on the relevance of their contents to what each test was expected to evaluate. In this case the relevance of the writing test was self-evident in the sense that the test called for the ability to express one's ideas in writing, which was exactly what the testees were expected to do. As to its degree of consistency in performing that task, the test was found to have a

reliability of .91 on the basis of homogeneity reliability analysis, or reader reliability of .93.

The relevance of the grammar test was similarly referred to its contents. They were entirely based on the description of grammar of BI in a source that could only be regarded as the most dependable reference available, as far as standard Indonesian grammar is concerned. And upon analysis of its results following the try-out test, the degree of reliability calculated using KR-20 formula, was found to be .87. A very similar figure (.88) was obtained for the cloze test, that was understood as measuring the global mastery of BI, that it did.

Test participants

The target group for the whom the test battery was constructed were the Indonesian educated people. These were represented by college graduates, who had finished a four-year S1 program at various colleges and universities. At the time of the test development the graduates were studying at graduate schools to qualify for S2 or S3 degrees. Altogether a total of 103 students of a variety of major studies took part in the test that was specially administered for the purpose at four out of the ten existing graduate schools. All data for the analysis of the test were obtained from these participants.

Besides the graduate students, however, ten college lecturers of BI were invited to do the same tests. Due to their formal qualification and experience in the teaching of language and BI, the lecturers were expected to supply with some data and information as point of reference and comparison for the data obtained from the regular participants, the graduate students.

Test results analysis

Writing test

Analysis of results of the writing test was based on a series of scoring done by two scorers working independently using the profile. The results showed among others that (1) the scores of the graduate students ranged from 34 to 80.5, equivalent to "poor" and "good" level of mastery; (2) almost half of the participants (45%) showed a "poor" writing ability, 50% "moderate", and only 5% "good"; (3) most of the lecturers, (9 of them or 90%) were "good" and one of them was even "very good".

Aside from the results of the test showing the unsatisfactory condition of the graduates' writing ability in BI, the test itself was apparently functioning

satisfactorily. Apart from a minor modification in the direction suggested later by members of the reference group, the writing test as it was administered and evaluated in the study, was an acceptable and feasible format to use as a valid and reliable test of writing of BI.

Grammar test

With a maximum score of 100 (2 points for each correct answer) the grammar test results showed a low mean score of 32.8 and scores ranging widely from 6 to 62. Even members of the reference group, surprisingly, did not really fare much better as one might expect. Although their range of scores was narrower (24 to 54), their average score was only 38 and the highest score was even lower compared to that of the graduate students (54:62). This was interpreted as an indication, as far as this grammar test was concerned, of the generally low level of mastery of grammar of BI among college graduates and lecturers alike.

Faced with a situation where no definite criterion was available, not even dependable points for comparison from the reference group, it was decided to rely on norm-referenced method of interpreting the scores for more definite qualifications of their degrees of mastery of grammar. At the application of principles and formulas of normal distribution using the available scores, the distribution of the participants' scores reflecting their grammar competence was found to be almost perfectly normal. Starting from the highest level of A ("very good") down to the lowest E ("very poor") the scores were normally distributed into A (2%), B (14%), C (66%), D (16%) and E (2%). For a group of college graduates, even those normal percentages did not present an ideal representation of their mastery of grammar of BI. A much better performance was to be expected.

The test itself, however, may somehow have been defective, and at least partly responsible for the low quality performance of those who took the test. Through more critical observation and examination of the test, where members of the reference group were involved, some possible deficiencies were identified, especially concerning the use of lengthy sentences for the test items. Perhaps not unrelated to that were the problems of difficulty level (too difficult) and time allocation (not enough). Steps to improve the grammar test should definitely give priority to those aspects.

Cloze test

Although in general the cloze test scores were higher than those of the grammar test, in many respects the results of the two tests showed similar features of the participants' performance. Out of a maximum score of 100 for the 50-item test, the average score was found to be 47. with a score range between 18 and 78. The scores of the reference group were better in terms of average score (55.6), and narrower range (between 48 and 64), although the highest score was again lower (64:78).

For similar reasons, interpretation of the cloze test scores was also based on norms rather than criterion that was not available. Upon simple calculations using the existing scores and norm-reference orientation, the levels of global language competence of the participants reflected in the cloze test were specified as A (2%), B (14%), C (74%), D (7%) and E (3%) respectively. Assuming that the test was of good quality, the performance of the college graduates was not really impressive.

Attempts to improve the cloze test may have to reconsider some aspects, especially concerning selection of the texts to ensure more neutral, less biased background orientation, and a more appropriate rate of deletion. These were some of the points observed later or obtained as part of the feedback from members of the reference group.

Test battery

Correlation analysis of the results of the three tests produced a higher degree of correlation between writing and cloze tests (.38) than writing and grammar (.27). A much higher degree of correlation was found between cloze and grammar tests (.59).

The correlation figure were interpreted as indicating (1) in relation to writing test as the main test to be developed in this study, the grammar test was more complementary than cloze test, due to its lower degree of correlation; (2) a substantial degree of similarity between cloze and grammar tests in terms of their capability in assessing language mastery, and therefore (3) a greater degree of mutual substitutability between the two, where the use of one may substitute for the other, implying that (4) cloze test was more dispensable as part of the test battery, leaving (5) writing test, the main test, and grammar as complementary components of the test battery.

Other results of the analysis showed that the correlation between the writing test as a whole and its grammar component as part of the profile, was very high (.94), much higher than correlation with the grammar test (.27). This might mean that despite the similar label, the

grammar component of the writing test and the grammar test may have actually reflected different grammatical points. In other words the kinds of grammatical aspect required for writing may not necessarily be the same as those in the grammar test that, after all, was based on the theoretical description and specification of grammar of B . If this was the case, then there was really a better rea n to consider the writing and grammar tests as complementary to each other. To include both of them in the initial development of Indonesian language test battery for college graduates would be a feasible endeavor. And with the writing test in an acceptable form and quality the next step would be improving the grammar test.

References

Badudu, Yus, Bahasa Indonesia: Anda Bertanya? Inilah Jawabnya, Bandung: Pustaka Prima, 1984.

Badudu, J.S., Cakrawala Bahasa Indonesia, Jakarta: Penerbit PT. Gramedia, 1985.

Departemen Pendidikan dan Kebudayaan, Tatabahasa Baku Bahasa Indonesia, Jakarta: Balai Pustaka, 1988.

Djiwandono, M. Soenardi, Pengembangan Tes Kemampuan Berbahasa Indonesia, IKIP Malang, 1990.

ENI (Ensiklopedi Nasional Indonesia), Jilid 1, Jakarta: PT Cipta Adi Pustaka, 1988.

Jacobs, Holly L., dkk, Testing ESL Composition: A Practical Approach, Rowley, Mass.: Newbury House Publisher, Inc., 1987.

PPPB (Pusat Pembinaan dan Pengembangan Bahasa) Hasil Perumusan Seminar, Seri Penyuluhan 3, 1977.

Sadtono, E., "Bahasa Indonesia Tertulis Golongan Terdidik (Sarjana) di Indonesia", dalam Pengajaran Bahasa dan Sastra, tahun ke-1 no.5, 1976.

Samsuri, "Mari Kita Laksanakan Cita-cita Sumpah Pemuda", dalam Kongres Bahasa Indonesia IV, Jakarta: PPPB, 1985.

Suharianto S., "Peranan Pengajaran Kemampuan Bahasa dalam Pengembangan Sastra Indonesia", dalam Pengajaran Bahasa dan Sastra, Th. III no.2, 1977.

Sunarjono Basuki dkk., Kemampuan Berbahasa Indonesia Siswa Kelas III SPG Negeri di Bali, Denpasar: Proyek Penelitian Bahasa dan Sastra Indonesia dan Daerah Bali, PPPB, Dep.PK, 1986.

Syafi'ie Imam, Analisis Kesalahan Berbahasa Indonesia dalam Menulis Mahasiswa Tiga IKIP di Jawa, disertasi, FPS IKIP Malang, 1984.